DOCUMENT RESUME

ED 306 552 CS 009 650

AUTHOR Powell, Janet L.
TITLE How Well Do Tests Measure Real Reading? ERIC Digest.
INSTITUTION ERIC Clearinghouse on Reading and Communication Skills, Bloomington, IN.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Jun 89
CONTRACT RI88062001
NOTE 4p.
PUB TYPE Information Analyses - ERIC Information Analysis Products (071)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Construct Validity; Elementary Education; Metacognition; Reading Comprehension; Reading Diagnosis; *Reading Processes; *Reading Research; *Reading Tests; *Testing Problems
IDENTIFIERS ERIC Digests; Reading Behavior

ABSTRACT

Despite a significant increase in test usage, numerous issues surrounding reading assessment remain unresolved. Construct validity--whether the test actually measures aspects of the behavior under consideration--is of particular importance if one is to rely on test scores to direct instruction, predict performance, or determine accountability. A slow but continually emerging trend to recognize reading as a thinking process has been at the core of the controversies over the validity of various forms of reading assessment. Virtually all methods of assessing reading are indirect, even those that claim to directly assess reading processes. The product of reading should, however, reflect the process the test-taker uses to generate the responses that produce a reading comprehension test score. A reader's awareness of thought processes involved in reading has recently been called "metacognition," and test designers are now including items that supposedly measure this. The as yet limited research on the metacognitive aspects of reading has indicated that multiple-choice tests and written retellings have construct validity. While the scores (products) of these tests may not reveal direct information about the processes students use to complete them, the tasks do appear to involve mental processes that have long been associated with reading. (MM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ERIC CLEARINGHOUSE ON READING AND COMMUNICATION SKILLS

June 1989 DIGEST

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED306552

How Well Do Tests Measure Real Reading?

by Janet L. Powell

Despite a significant increase in test usage across the country, numerous issues surrounding the testing of reading remain unresolved. (See Johnston, 1986.) How validly it reflects what people actually do when they read is the most important consideration of any reading test. Construct validity – whether the test actually measures aspects of the behavior under consideration – is of particular importance if one is to rely on test scores to direct instruction, predict performance, or determine accountability. In 1917, Thorndike (see 1971 reprint), who defined reading as reasoning, helped promote the examination of reading as a cognitive process as thought guided by printed symbols (Farr and Roser, 1979).

Are we measuring process?

This slowly but continually emerging trend to recognize reading as a thinking process has been at the core of the controversies over the validity of various forms of reading assessment. Many critics of reading tests claim that most current approaches to the assessment of reading comprehension remain – as they have always been – measures of reading comprehension as a product of a reader's interaction with a text. Unable to assess the processes involved in comprehension, the tests measure comprehension as required responses that are the products of reading (Johnston, 1983).

Virtually all methods of assessing reading are indirect, even those that *claim* to directly assess reading *processes*. We cannot actually see the processes involved; we can only infer how a reader has comprehended. Therefore, all scores or data produced by tests of reading are indirect measures of the reading process.

The product of reading should, however, reflect the process the test-taker uses to generate the responses that produce a reading comprehension test score. That is to say that one ought to be able to assume that differences in test scores across test-takers and testing instances will reflect differences in the processes used to read the test passage and to respond as directed. How *directly* the two relate has never been determined; nor do we know how effectively test results can inform and direct the teaching of reading behaviors – even when those behaviors appear to be very similar to those that produce the test product. How well tests that do not emphasize or examine product might direct instruction that purports to develop process is a matter even less well understood.

Farr (1986) states that "the manuals of most standardized tests make very explicit the fact that the test will not provide information about a pupil's reading *processes*, but only information about the *product* of reading." However, he continues by saying that "...one could argue that the product – or score – isn't valid if a pupil doesn't use the actual processes of reading in determining the answers." The validity question that surrounds the tests thus seems to be whether or not taking the test *appears* to change the processes involved in comprehension and to solicit significantly atypical reading processes.

Metacognition focuses on process

A reader's awareness of thought processes involved in reading has recently come to be known as *metacognition*, and test designers are now including items that supposedly measure this (Aronson and Farr, 1988). The general knowledge of the reader guides him or her in monitoring comprehension processes through the selection and implementation of specific strategies to achieve some predetermined goal or purpose for reading. The chief idea involved in metacognition is that learners must actively monitor their use of thinking processes – that they must be aware of how they are processing information – and that they can then regulate them according to the purpose for reading.

CS009650

The interest in metacognition among reading educators has led to an exploration of procedures to collect data on thinking processes. Data collection on mental processes has become known as *introspective data—concurrent* and *retrospective* verbal reports. Concurrent verbal reports are collected *as the subject is engaged in the reading task* .These types of reports have been criticized for interfering with the normal processes of reading (Nisbett and Wilson, 1977; Garner, 1982). Retrospective verbal reports are collected after the subject has completed the reading task. These types of reports have been criticized because subjects may forget or inaccurately recall the mental processes they employed while completing the task (Afflerbach and Johnston, 1984).

There are differences of opinion as to the validity and reliability of verbal report data in general. However, many prominent researchers agree that verbal reports, when they are elicited with care and interpreted with full understanding of the circumstances under which they were obtained, are valuable and thoroughly reliable sources of information about cognitive processes (Afflerbach and Johnston, 1984).

Verbal reports may reveal reading processes

Research that focuses on the metacognitive aspects of reading while taking a reading test comprise only a very small portion of the literature. At least three studies, however, have used verbal reports to investigate reading processes as subjects are engaged in taking reading compre- hension tests. Using concurrent verbal reports, Wingenbach (1984) examined the comprehension processes employed by twenty gifted readers in grades 4 through 7 to identify the metacognitive strategies they employed as they read the *Iowa Test of Basic Skills,* a multiple-choice standardized reading test.

Wingenbach found that subjects reported using a variety of reading strategies to comprehend the text and to answer the questions. The strategies included using context clues, rereading, inferencing, personal identification with the text, and imagery. Wingenbach did not use as a comparison any other text types, making it impossible to determine whether or not the subject's mental processing was different on the test than on any other reading task.

Alvermann and Ratekin (1982) conducted a study with 98 "average" seventh-grade and eighth-grade subjects. The subjects completed a multiple-choice test and an essay test. Only retrospective reports were collected. Results of an analysis of the verbal protocols revealed that 55 subjects reported using only one reading strategy, while 30 reported using two or more. Thirteen subjects were unable to recall any specific strategy. In the report, Alvermann and Ratekin elaborate only on the statistically significant differences in strategies. They found that subjects who read to respond on an essay test "reread" more frequently than students who read the same passage knowing they will respond to multiple-choice items. In addition, subjects who read to complete an essay test reported using multiple strategies nearly twice as often as students who read for a multiple-choice test.

Other differences that were not statistically significant, may be important nevertheless. An examination of a chart representing the frequency of reported strategies shows that students read for details twice as often in the multiple-choice test as they did in the essay test. There were four reports of imaging (forming a picture of the text) in the essay test compared to one in the multiple-choice test. Subjects made a personal connection with the text an average of seven times when taking the multiple-choice test but only three during the essay test.

The use of only retrospective verbal reports severely limits the conclusions made by the researchers. When retrospection alone is used, the chances that the subjects forgot the mental processes they employed are greatly enhanced. In addition, the differences found may have been due to individual or group differences rather than task-related differences. There is little information in the report to support that the two groups were equivalent.

Powell (1988) conducted a study with nine proficient sixth-grade readers. All the subjects were observed, and they provided concurrent verbal reports as they were engaged in multiple-choice tests, cloze tests, written retellings, and a non-assessed reading task. The subjects gave retrospective verbal reports afterward. Twenty-one reading processes were identified from the verbal reports. The overall conclusions of this investigation indicated that the reading processes did differ as subjects were engaged in each of the tasks. The task which elicited behavior the most different from the other three was the cloze test. Subjects reported rereading and using context clues a great deal more

on this task than on any of the others. They tied prior knowledge to the text and paraphrased the text a great deal less than in performing the other reading tasks.

The multiple-choice test and the written retellings,on the other hand, were very similar to each other and to the nonassessed reading task. The subjects reported tying prior knowledge in with the text, visualizing what was happening in the text, and paraphrasing the text almost with equal frequency across all three tasks. Therefore, within the limitations of the Powell study, it can be concluded that multiple-choice tests and written retellings had construct validity. While the scores (products) of these tests may not reveal direct information on the processes students use to complete them, the tasks do appear to involve mental processes that have long been associated with reading.

References

Afflerbach, Peter, and Johnston, Peter. "On the use of verbal reports in reading research," *Journal of Reading Behavior, 16* (4), 1984, pp. 307-322.

Alverman, Donne, and Ratekin, Ned. "Metacognitive knowledge about reading proficiency: Its relation to study strategies and task demands." Paper presented at the Annual Meeting of the American Educational Research Association, 1982.

Aronson, Edith, and Farr, Roger. "Issues in assessment," *Journal of Reading*, 32 (2), November 1988, pp. 174-177.

Farr, Roger. "A response to Hoopfer and Hunsberger," *Forum in Reading and Language Education*, October 1986, pp. 129-133.

Farr, Roger, and Roser, Nancy. *Teaching a child to read.* New York: Harcourt Brace Javanovich, 1977.

Garner, Ruth. "Verbal-report data on reading strategies," *Journal of Reading Behavior, 14* (2), 1982, pp. 159-167.

Johnston, Peter. *Reading comprehension assessment: A cognitive basis.* Newark, Delaware: International Reading Association, 1983, 102 pp. [ED 226 338]

Johnston, Peter. "Assessing the process and the process of assessment in the language arts." In James Squire (Ed.), *The dynamics of language learning: Research in the language arts.* Urbana, Illinois: National Council of Teachers of English, 1986, 420 pp. [ED 280 080]

Meichenbaum, D.; Burland, S.; Gruson, L.; and Cameron, R. Metacognitive assessment. In S. R. Yussen (Ed.), *The growth of reflection in children.* Orlando: Academic Press, 1985.

Nisbett, Richard, and Wilson, Timothy. (1977). "Telling more than we can know: Verbal reports on mental processes," *Psychological Review, 84* (3), 1977, pp. 231-259.

Powell, Janet L. "An examination of comprehension processes used by readers as they engage in different forms of assessment. "Unpublished Doctoral Dissertation, Indiana University, 1988. 196 pp. [ED 298 449]

Thorndike, Edward."Reading as reasoning: A study of mistakes in paragraph reading," *Reading Research Quarterly, 6* (4), 1971 [reprint of original 1917 article], pp. 425-434.

Wingenbach, Nancy (1984). "The gifted reader: Metacognition and comprehension strategies." Paper presented at the 3rd Annual Meeting of the National Council of Teachers of English, Spring Conference, 1984. 29pp. [ED 243 093]

ERIC Clearinghouse on
Reading and Communication Skills
Indiana University
Smith Research Center
2805 East Tenth Street, Suite 150
Bloomington, IN 47408
(812) 855-5847

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract no. RI88062001. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions, however, do not necessarily represent the official view or opinions of the Office of Educational Research and Improvement.